

# Reed-Muller Codes Achieve Capacity on the Binary Erasure Channel under MAP Decoding

Shrinivas Kudekar<sup>‡</sup>, Marco Mondelli<sup>\*</sup>, Eren Şaşoğlu<sup>†</sup>, Rüdiger Urbanke<sup>\*</sup>,

<sup>\*</sup>School of Computer and Communication Sciences, EPFL, Switzerland

Emails: {marco.mondelli, ruediger.urbanke}@epfl.ch

<sup>†</sup>UC Berkeley

Email: eren.sasoglu@gmail.com

<sup>‡</sup>Qualcomm Research, New Jersey, USA

Email: skudekar@qti.qualcomm.com

**Abstract**—We show that Reed-Muller codes achieve capacity under maximum a posteriori bit decoding for transmission over the binary erasure channel for all rates  $0 < R < 1$ . The proof is generic and applies to other codes with sufficient amount of symmetry as well. The main idea is to combine the following observations: (i) monotone functions experience a sharp threshold behavior, (ii) the extrinsic information transfer (EXIT) functions are monotone, (iii) Reed-Muller codes are 2-transitive and thus the EXIT functions associated with their codeword bits are all equal, and (iv) therefore the Area Theorem for the average EXIT functions implies that RM codes' threshold is at channel capacity.

**Keywords**—RM codes, MAP decoding, capacity-achieving codes, BEC, EXIT function

## I. INTRODUCTION

Reed-Muller (RM) codes [1]–[4] are among the oldest codes in existence, and due to their many desirable properties, are also among the most widely studied. In recent years there has been renewed interest in RM codes, partly due to the invention of capacity-achieving polar codes [5], which are closely related to RM codes. For a performance comparison between polar and RM codes, see [6], [7]. Simulations and analytical results suggest that RM codes do not perform well under successive and iterative decoding, but they outperform polar codes under maximum a posteriori (MAP) decoding [5], [8]. Nevertheless, it is not known whether RM codes themselves are capacity-achieving except for rates approaching 0 and 1 over the binary erasure channel (BEC) and the binary symmetric channel (BSC) [9].

In this paper, we show that RM codes indeed achieve the capacity for transmission over the BEC for any rate  $R \in (0, 1)$ . The same result was shown independently by Kumar and Pfister [10] using essentially the same approach.

## II. MAIN RESULT

Let  $RM(n, r)$  denote the Reed-Muller (RM) code of block length  $N = 2^n$  and order  $r$ , see [3]. This is a linear code of rate  $R = \frac{1}{N} \sum_{i=0}^r \binom{n}{i}$  and minimum distance  $d = 2^{n-r}$ , generated by all rows of weight at least  $2^{n-r}$  of the Hadamard matrix  $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\otimes n}$ , where  $\otimes$  denotes the Kronecker product. Let  $[N] = \{1, \dots, N\}$  denote the index set of codeword bits. For  $i \in [N]$ , let  $x_i$  denote the  $i$ th component of a vector  $x$ , and let  $x_{\sim i}$  denote the vector containing all components except

$x_i$ . For  $x, y \in \{0, 1\}^N$ , we write  $x \prec y$  if  $y$  dominates  $x$  component-wise, i.e. if  $x_i \leq y_i$  for all  $i \in [N]$ .

Let  $BEC(\epsilon)$  denote the binary erasure channel with erasure probability  $\epsilon$ . Recall that this channel has capacity  $1 - \epsilon$  bits/channel use. In what follows, we will fix a rate  $R$  for a sequence of RM codes and show that the bit error probability of the code sequence vanishes for all BECs with capacity strictly larger than  $R$ , i.e., erasure probability strictly smaller than  $1 - R$ .

**Theorem 1 (RM Codes Achieve Capacity on the BEC):**

Consider a sequence of  $RM(n, r_n)$  codes of increasing  $n$  and rate  $R_n$  converging to  $R$ ,  $0 < R < 1$ . For any  $0 \leq \epsilon < 1 - R$  and any  $\delta > 0$  there exists an  $n_0$  such that for all  $n > n_0$  the bit error probability of  $RM(n, r_n)$  is bounded above by  $\delta$  under bit-MAP decoding.

The only property of RM codes that has a bearing on the following proof of Theorem 1 is that these codes exhibit a high degree of symmetry, and in particular, that they are invariant under a 2-transitive group of permutations on the coordinates of the code [3], [11], [12]. In fact, this proof also shows that all 2-transitive sequences of codes are capacity-achieving. We will return to this point in Section III.

**Lemma 1 (RM Codes Are 2-Transitive):** For any  $a, b, c$ , and  $d \in [N]$  s.t.  $a \neq b$  and  $c \neq d$ , there exists a permutation  $\pi : [N] \rightarrow [N]$  such that

- (i)  $\pi(a) = c, \pi(b) = d$ , and
- (ii)  $RM(n, r)$  is closed under the permutation of its codeword bits according to  $\pi$ . That is,

$$\begin{aligned} (x_1, \dots, x_N) &\in RM(n, r) \\ &\Updownarrow \\ (x_{\pi(1)}, \dots, x_{\pi(N)}) &\in RM(n, r). \end{aligned} \tag{1}$$

The 2-transitivity of the code implies many symmetries that will be critical in the proof, which we outline here. We will be interested in MAP decoding of the  $i$ th codebit  $x_i$  from observations  $y_{\sim i}$ , that is, all channel outputs except  $y_i$ . The error probability of the  $i$ th such decoder for transmission over a  $BEC(\epsilon)$  is called the  $i$ th EXIT function [13, Lemma 3.74], which we denote by  $h_i(\epsilon)$ . We will see that all  $N$  EXIT functions of an RM code (and of any 2-transitive code) are identical, and also that erasure patterns that lead to decoding

errors under this decoder exhibit a high degree of symmetry. These symmetries will imply that the EXIT functions have a sharp threshold behavior, i.e., the bit error probability is very small below a threshold, and very large above. A final and crucial benefit of considering this suboptimal decoder and EXIT functions instead of the optimal block-MAP decoder is the well-known *Area Theorem* [13]–[16], which will allow us to show that the threshold is at channel capacity and conclude the proof.

Recall the basic definition of an EXIT function [13, Lemma 3.74] and its relation to bit-MAP decoding.

*Definition 1 (EXIT Function):* Let  $\mathcal{C}[N, K]$  be a binary linear code of rate  $R = K/N$  and let  $X$  be chosen with uniform probability from  $\mathcal{C}[N, K]$ . Let  $Y$  denote the result of letting  $X$  be transmitted over a BEC( $\epsilon$ ). The EXIT function  $h_i(\epsilon)$  associated with the  $i$ th bit of  $\mathcal{C}$  is defined as

$$h_i(\epsilon) = H(X_i | Y_{\sim i}). \quad (2)$$

*Lemma 2 (EXIT Function and Bit-MAP Decoding):* Let  $\mathcal{C}[N, K]$  be a binary linear code and let  $\hat{x}^{\text{MAP}}(y_{\sim i})$  denote the MAP estimator of the  $i$ th code bit given the observation  $y_{\sim i}$ . Then,

$$h_i(\epsilon) = \mathbb{P}(\hat{x}^{\text{MAP}}(Y_{\sim i}) = ?). \quad (3)$$

The most relevant property of EXIT functions for our purpose is the Area Theorem, see [13]–[16].

*Lemma 3 (Area Theorem):* Let  $\mathcal{C}[N, K]$  be a binary linear code, and let  $h(\epsilon) = \frac{1}{N} \sum_{i=1}^{N-1} h_i(\epsilon)$  be the *average* EXIT function. Then,

$$\int_0^\epsilon h(x) dx = \frac{1}{N} H(X | Y),$$

where  $H(X | Y)$  is the conditional entropy of the codeword  $X$  given the observation  $Y$  at the receiver. In particular,

$$\int_0^1 h(x) dx = R = \frac{K}{N}.$$

We now show that the erasure patterns that lead to decoding failures are monotone and symmetric. Recall that the decoding of each bit relies only on  $N - 1$  received bits. We will denote each erasure pattern by a binary vector of length  $N - 1$ , where a 1 denotes an erasure and a 0 denotes a non-erasure. We first characterize the set  $\Omega_i$  that leads to a decoding failure for bit  $i$ .

*Definition 2 ( $\Omega_i$ ):* Given a binary linear code  $\mathcal{C}[N, K]$ , let  $\Omega_i$  be the set that consists of all  $\omega \in \{0, 1\}^{N-1}$  for which there exists  $c \in \mathcal{C}$  such that  $c_i = 1$  and  $c_{\sim i} \prec \omega$ .

*Lemma 4 ( $\Omega_i$  Encodes  $h_i(\epsilon)$ ):* Let  $\omega \in \{0, 1\}^{N-1}$  be the erasure pattern on the received bits  $y_{\sim i}$ . Then the  $i$ th bit-MAP decoder fails if and only if  $\omega \in \Omega_i$ . Consequently, if  $\mu_\epsilon(\cdot)$  is the measure on  $\{0, 1\}^{N-1}$  that puts weight  $\epsilon^w(1 - \epsilon)^{N-1-w}$  on a point of Hamming weight  $w$ , then

$$h_i(\epsilon) = \mu_\epsilon(\Omega_i).$$

That is,  $\Omega_i$  “encodes” the EXIT function of the  $i$ th position.

*Proof:* Since the code is linear and the channel is symmetric and memoryless, we can assume that the all-zero codeword was transmitted. Given an erasure pattern  $\omega$ , let  $\mathcal{C}'$  denote the

set of all codewords  $c$  that are compatible with the observation  $y_{\sim i}$ , i.e., all codewords for which  $c_{\sim i} \prec \omega$ . Note that since the code is linear, so is  $\mathcal{C}'$ . This implies that if there exists a  $c \in \mathcal{C}'$  with  $c_i = 1$ , then half of all codewords in  $\mathcal{C}'$  have a 0 at position  $i$ , and the other half have a 1, and thus the bit-MAP decoder fails to decode bit  $i$ . On the other hand, if there is no  $c \in \mathcal{C}'$  with  $c_i = 1$ , then all compatible codewords have a 0 at position  $i$ , and thus the bit-MAP decoder succeeds. That is,  $\Omega_i$  is the set of all erasure patterns s.t. the bit-MAP decoder cannot decide on position  $i$  given the observation  $y_{\sim i}$ . The claim that  $h_i(\epsilon) = \mu_\epsilon(\Omega_i)$  follows immediately, since the memorylessness of the channel implies that an erasure pattern  $\omega$  occurs with probability  $\mu_\epsilon(\omega)$ . ■

*Lemma 5 ( $\Omega_i$  is Monotone):* If  $\omega \in \Omega_i$  and  $\omega \prec \omega'$ , then  $\omega' \in \Omega_i$ .

*Proof:* If  $\omega \in \Omega_i$ , then there exists a codeword  $c$  so that  $c_i = 1$  and  $c_{\sim i} \prec \omega$ . Since by assumption  $\omega \prec \omega'$ , it follows that  $c_{\sim i} \prec \omega'$ , which implies  $\omega' \in \Omega_i$ . ■

*Lemma 6 ( $\Omega_i$  is Symmetric):* If  $\mathcal{C}[N, K]$  is a 2-transitive binary linear code, then  $\Omega_i$  is invariant under a 1-transitive group of permutations for any  $i \in [N]$ . Following [17], we say that  $\Omega_i$  is *symmetric*.

*Proof:* Since  $\mathcal{C}$  is 2-transitive, for any  $j_1, j_2 \in [N] \setminus \{i\}$ , there exists a permutation  $\pi : [N] \rightarrow [N]$  so that

- $\pi(i) = i$ ,
- $\pi(j_1) = j_2$ ,
- $(c_{\pi(1)}, \dots, c_{\pi(N)}) \in \mathcal{C}$  for any  $(c_1, \dots, c_N) \in \mathcal{C}$ .

Let  $S_1 : [N - 1] \rightarrow [N] \setminus \{i\}$  be defined as  $S_1(k) = k$  for  $k \in \{1, \dots, i - 1\}$  and  $S_1(k) = k + 1$  for  $k \in \{i, \dots, N - 1\}$ . Let  $S_2 : [N] \setminus \{i\} \rightarrow [N - 1]$  be defined as  $S_2(k) = k$  for  $k \in \{1, \dots, i - 1\}$  and  $S_2(k) = k - 1$  for  $k \in \{i + 1, \dots, N\}$ . Consider the permutation  $\hat{\pi} : [N - 1] \rightarrow [N - 1]$  defined as  $\hat{\pi}(k) = S_2(\pi(S_1(k)))$ . Note that, by changing the choice of  $j_1$  and  $j_2$ , we generate the 1-transitive group of permutations on  $[N - 1]$ . It then suffices to show that if  $\omega = (\omega_1, \dots, \omega_{N-1}) \in \Omega_i$ , then  $(\omega_{\hat{\pi}(1)}, \dots, \omega_{\hat{\pi}(N-1)}) \in \Omega_i$ .

Recall that  $\omega \in \Omega_i$  if there exists a codeword  $c = (c_1, \dots, c_N) \in \mathcal{C}$  so that  $c_i = 1$  and  $c_{\sim i} \prec \omega$ . By construction of  $\pi$ , we have that  $(c_{\pi(1)}, \dots, c_{\pi(N)}) \in \mathcal{C}$  and, in addition,  $c_{\pi(i)} = c_i = 1$ . By construction of  $\hat{\pi}$ ,  $(c_{\pi(1)}, \dots, c_{\pi(i-1)}, c_{\pi(i+1)}, \dots, c_{\pi(N)}) \prec (\omega_{\hat{\pi}(1)}, \dots, \omega_{\hat{\pi}(N-1)})$ . As a result,  $(\omega_{\hat{\pi}(1)}, \dots, \omega_{\hat{\pi}(N-1)}) \in \Omega_i$  and the proof is complete. ■

We now show that all EXIT functions of a 2-transitive code are identical.

*Lemma 7 ( $h_i$  is Independent of  $i$ ):* If  $\mathcal{C}[N, K]$  is a 2-transitive binary linear code, then  $h_i(\epsilon) = h_j(\epsilon)$  for all  $i, j \in [N]$ . That is,  $h_i(\epsilon)$  is independent of  $i$ .

*Proof:* Since  $\mathcal{C}$  is 2-transitive, there exists a permutation  $\pi : [N] \rightarrow [N]$  so that

- $\pi(i) = j$ ,
- $(c_{\pi(1)}, \dots, c_{\pi(N)}) \in \mathcal{C}$  for any  $(c_1, \dots, c_N) \in \mathcal{C}$ .

Let  $S_i : [N - 1] \rightarrow [N] \setminus \{i\}$  be defined as  $S_i(k) = k$  for  $k \in \{1, \dots, i - 1\}$  and  $S_i(k) = k + 1$  for  $k \in \{i, \dots, N - 1\}$ .

Let  $S_j : [N] \setminus \{j\} \rightarrow [N-1]$  be defined as  $S_j(k) = k$  for  $k \in \{1, \dots, j-1\}$  and  $S_j(k) = k-1$  for  $k \in \{j+1, \dots, N\}$ . Consider the permutation  $\hat{\pi} : [N-1] \rightarrow [N-1]$  defined as  $\hat{\pi}(k) = S_j(\pi(S_i(k)))$ .

Pick  $\omega \in \Omega_j$ . Then, there exists a codeword  $c$  so that  $c_j = 1$  and  $c_{\sim j} \prec \omega$ . By construction of  $\pi$ , we have that  $(c_{\pi(1)}, \dots, c_{\pi(N)}) \in \mathcal{C}$  and, in addition,  $c_{\pi(i)} = c_j = 1$ . By construction of  $\hat{\pi}$ ,  $(c_{\pi(1)}, \dots, c_{\pi(i-1)}, c_{\pi(i+1)}, \dots, c_{\pi(N)}) \prec (\omega_{\hat{\pi}(1)}, \dots, \omega_{\hat{\pi}(N-1)})$ . As a result,  $(\omega_{\hat{\pi}(1)}, \dots, \omega_{\hat{\pi}(N-1)}) \in \Omega_i$ .

With an abuse of notation, let us define

$$\hat{\pi}(\Omega_j) = \{(\omega_{\hat{\pi}(1)}, \dots, \omega_{\hat{\pi}(N-1)}) : \omega \in \Omega_j\}.$$

Then, the previous argument implies that  $\hat{\pi}(\Omega_j) \subseteq \Omega_i$ .

It is clear that, if  $\omega \neq \omega'$ , then  $(\omega_{\hat{\pi}(1)}, \dots, \omega_{\hat{\pi}(N-1)}) \neq (\omega'_{\hat{\pi}(1)}, \dots, \omega'_{\hat{\pi}(N-1)})$ . Indeed, if  $\omega \neq \omega'$ , then there exists an index  $k$  s.t.  $\omega_k \neq \omega'_k$  and, therefore,  $\omega_{\hat{\pi}(k)} \neq \omega'_{\hat{\pi}(k)}$ . In addition, the permutation  $\hat{\pi}$  leaves the weight of  $\omega$  unchanged. As a result, we have

$$h_j(\epsilon) = \mu_\epsilon(\Omega_j) \stackrel{(a)}{=} \mu_\epsilon(\hat{\pi}(\Omega_j)) \stackrel{(b)}{\leq} \mu_\epsilon(\Omega_i) = h_i(\epsilon), \quad (4)$$

where (a) comes from the fact that the channel acts independently and identically on each component, and (b) follows from  $\hat{\pi}(\Omega_j) \subseteq \Omega_i$ . By repeating the same argument with the indices  $i$  and  $j$  exchanged, we obtain opposite inequality and, therefore, the thesis follows. ■

We recall here the main ingredient for our proof, due to Friedgut and Kalai. We note that Tillich and Zémor applied the following theorem in [18] to show that *every* sequence of linear codes of increasing Hamming distance has a sharp threshold under block-MAP decoding for transmission over the BEC and the BSC.

*Theorem 2 (Sharp Threshold – [17]):* Let  $\Omega \in \{0, 1\}^N$  be a symmetric monotone set, where symmetry and monotonicity are defined as in Lemma 5 and 6, respectively. If  $\mu_\epsilon(\Omega) > \delta$ , then  $\mu_{\bar{\epsilon}}(\Omega) > 1 - \delta$  for  $\bar{\epsilon} = \epsilon + c \frac{\log(\frac{1}{2\delta})}{\log(N)}$ , where  $c$  is an absolute constant.

*Proof of Theorem 1:* Consider a sequence of codes  $\text{RM}(n, r_n)$  with rates converging to  $R$ . That is, the  $n$ th code in the sequence has a rate  $R_n \leq R + \delta_n$ , where  $\delta_n \rightarrow 0$  as  $n \rightarrow \infty$ .

Lemma 7 implies that  $h_i(\epsilon)$  is independent of  $i$ , and, thus, it is equal to the average EXIT function  $h(\epsilon)$ . Therefore, by Lemma 3 we have

$$\int_0^1 h_i(\epsilon) d\epsilon = R_n \leq R + \delta_n.$$

Consider the set  $\Omega_i$  defined in Definition 2 that encodes  $h_i(\epsilon)$ . By Lemmas 5 and 6,  $\Omega_i$  is monotone and symmetric. Therefore, from Lemma 2 we have that if  $h_i(\bar{\epsilon}) = 1 - \delta$ , then  $h_i(\underline{\epsilon}) \leq \delta$  for  $\bar{\epsilon} = \underline{\epsilon} + c \frac{\log(\frac{1}{2\delta})}{\log(N-1)}$ , where  $c$  is an absolute constant.

Now, the function  $h_i(\epsilon)$  is increasing, and therefore by Lemma 2, the error probability of the  $i$ th bit-MAP decoder

is upper bounded by  $\delta$  for all  $i \in [N]$  and  $\epsilon \leq \underline{\epsilon}$ . In order to conclude the proof, it suffices to show that  $\underline{\epsilon}$  is close to  $1 - R$ . Note that by definition of  $\bar{\epsilon}$ , the area under  $h_i(\epsilon)$  is at least equal to

$$(1 - \bar{\epsilon})(1 - \delta) \geq 1 - \bar{\epsilon} - \delta = 1 - \underline{\epsilon} - c \frac{\log(\frac{1}{2\delta})}{\log(N-1)} - \delta.$$

On the other hand, this area is at most equal to  $R + \delta_n$ . Combining these two inequalities we obtain

$$\underline{\epsilon} \geq 1 - R - \delta - \delta_n - c \frac{\log(\frac{1}{2\delta})}{\log(N-1)}. \quad (5)$$

We see that  $\underline{\epsilon}$  can be made arbitrarily close to  $1 - R$  by picking  $\delta$  sufficiently small and  $N$  sufficiently large. That is, the bit error probability can be made arbitrarily small at rates arbitrarily close to  $1 - R$ . ■

### III. GENERALIZATIONS AND DISCUSSION

As mentioned above, the foregoing arguments hold for all 2-transitive codes, and not just RM codes. That is, all such codes are capacity achieving over the BEC under bit-MAP decoding. This includes, for example, the class of extended BCH codes ([3, Chapter 8.5, Theorem 16]).

RM codes are only one possible family of codes that can be derived from the Hadamard matrix. It is reasonable to assume that any subset of generators of sufficient weight from the Hadamard matrix will produce good codes. It would be interesting to see if such a statement can be proved. Clearly, the symmetries of RM codes that are used here will not be present in general.

Perhaps of even greater interest is whether RM codes achieve capacity on general binary-input memoryless output-symmetric channels and if the above technique can be extended. Note that it suffices to prove that RM codes achieve capacity *for the BSC* since (up to a small factor) the BSC is the worst channel, see [19, pp. 87–89]. Most of the notions that we used here for the BEC have a straightforward generalization (e.g., GEXIT functions replace EXIT functions) or need no generalization (2-transitivity). However, it is currently unclear if the GEXIT function can be encoded in terms of a monotone function. It is likely that different techniques will be needed to show sharp thresholds for GEXIT functions.

One of the main motivations for studying RM codes is their superior empirical performance (over the BEC) compared with the capacity-achieving polar codes. By far the most important practical question is whether this promised performance can be harnessed at low complexities.

### ACKNOWLEDGEMENT

This work was done while the authors were visiting the Simons Institute for the Theory of Computing, UC Berkeley. We would like to thank the Institute for providing us with a fruitful work environment. We further gratefully acknowledge discussions with Tom Richardson, Hamed Hassani, and in particular with Henry Pfister.

## REFERENCES

- [1] D. E. Muller, "Application of boolean algebra to switching circuit design and to error detection," *IRE Trans. Electronic Computers*, vol. EC-3, no. 3, pp. 6–12, 1954.
- [2] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *IRE Trans. Electronic Computers*, vol. 4, no. 4, pp. 38–49, 1954.
- [3] F. J. MacWilliams and N. J. A. Sloane, *Theory of Error-Correcting Codes*. NorthHolland, 1977.
- [4] I. Dumer, "Recursive decoding and its performance for low-rate Reed-Muller codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 811–823, May 2004.
- [5] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.
- [6] E. Arıkan, "A survey of Reed-Muller codes from polar coding perspective," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Jan. 2010, pp. 1–5.
- [7] —, "A performance comparison of polar codes and Reed-Muller codes," *IEEE Commun. Lett.*, vol. 12, no. 6, pp. 447–449, June 2008.
- [8] N. Hussami, S. B. Korada, and R. Urbanke, "Performance of polar codes for channel and source coding," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, July 2009, pp. 1488–1492.
- [9] E. Abbe, A. Shpilka, and A. Wigderson, "Reed-Muller codes for random erasures and errors," in *STOC*, 2015.
- [10] S. Kumar and H. Pfister, "Reed-Muller codes achieve capacity on erasure channels," May 2015, [Online]. Available: <http://arxiv.org/abs/1505.05123>.
- [11] T. Kasami, L. Shu, and W. Peterson, "New generalizations of the Reed-Muller codes-I: Primitive codes," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, pp. 189–199, Mar. 1968.
- [12] T. Berger and P. Charpin, "The automorphism group of generalized Reed-Muller codes," *Discrete Mathematics*, vol. 117, pp. 1–17, 1993.
- [13] T. Richardson and R. Urbanke, *Modern Coding Theory*. New York, NY, USA: Cambridge University Press, 2008.
- [14] A. Ashikhmin, G. Kramer, and S. ten Brink, "Code rate and the area under extrinsic information transfer curves," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 2002, p. 115.
- [15] —, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov 2004.
- [16] C. Méasson, A. Montanari, and R. Urbanke, "Maxwell's construction: the hidden bridge between maximum-likelihood and iterative decoding," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5277 – 5307, Dec. 2008.
- [17] E. Friedgut and G. Kalai, "Every monotone graph property has a sharp threshold," *Proc. Amer. Math. Soc.*, vol. 124, pp. 2993–3002, 1996.
- [18] J.-P. Tillich and G. Zémor, "Discrete isoperimetric inequalities and the probability of a decoding error," *Combinatorics, Probability and Computing*, vol. 9, pp. 465–479, 2000. [Online]. Available: [http://journals.cambridge.org/article\\_S0963548300004466](http://journals.cambridge.org/article_S0963548300004466)
- [19] E. Şaşıoğlu, "Polar coding theorems for discrete systems," Ph.D. dissertation, EPFL, 2011.